Assignment 35.1

Problem Statement:

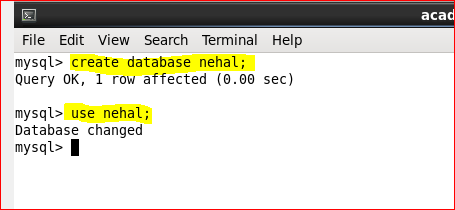
● Create a employee table in Mysql and columns as Emp\_id, Emp\_name, Dept\_name (Hadoop Developer), Emp\_sal.

● Import the employee table contents into the HDFS directory using Sqoop.

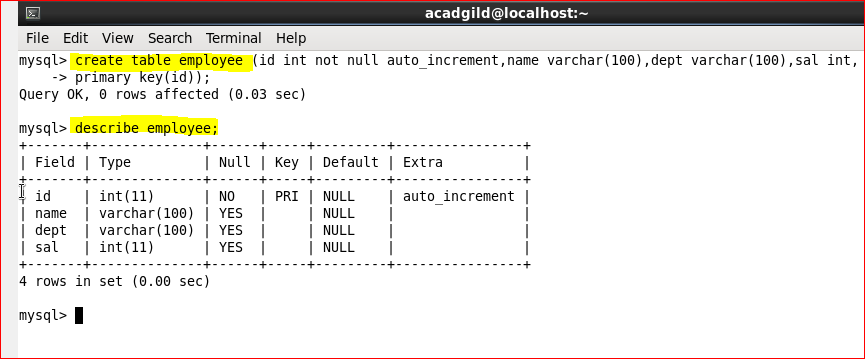
● Create a employee2 table in Mysql and Export employee details file from HDFS directory to Mysql table employee2 using Sqoop.

● Explain the procedures performed, Share the screenshots of commands and results for the same.

Create a database by using the below command:

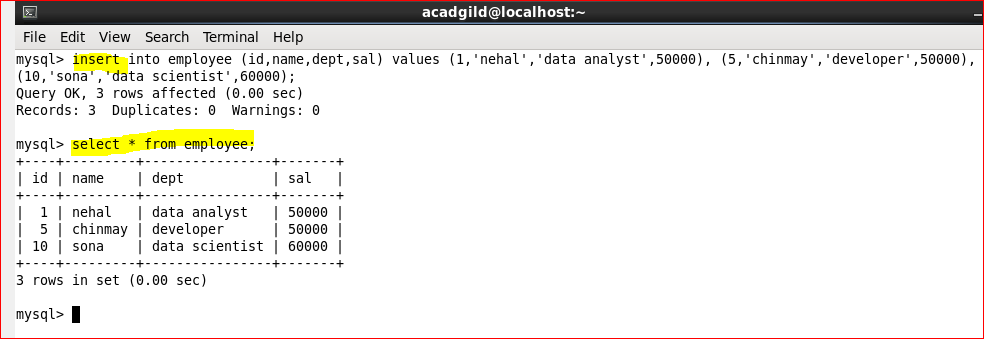


Create a table **employee**by using the below command:



A table has been created with name **employee**and with the columns **id, name, dept and sal.**

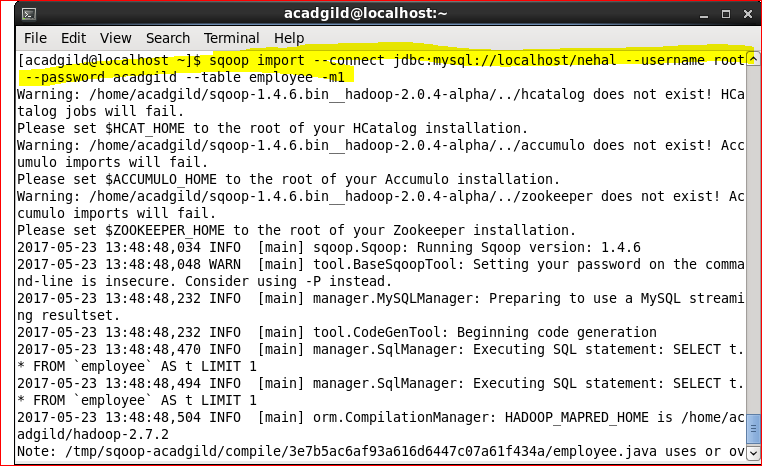
Let’s now try inserting some sample data into the created table by using the below command:



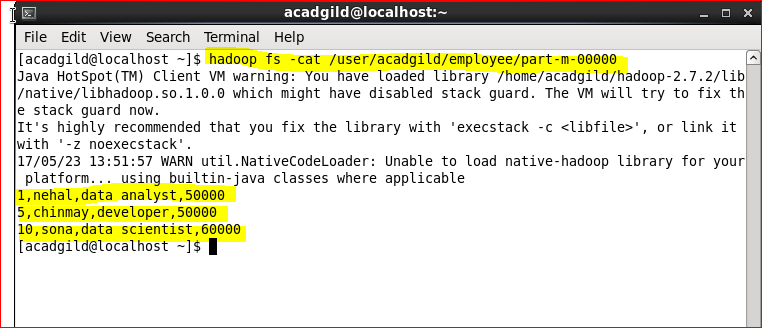
We have successfully created a table in MySQL, and we will now import the same into HDFS by using Sqoop.

Here we are connecting to MySQL through JDBC connectors and using the database **nehal.**Here it is necessary to specify the **MySQL‘s username and password**and the **table name.**

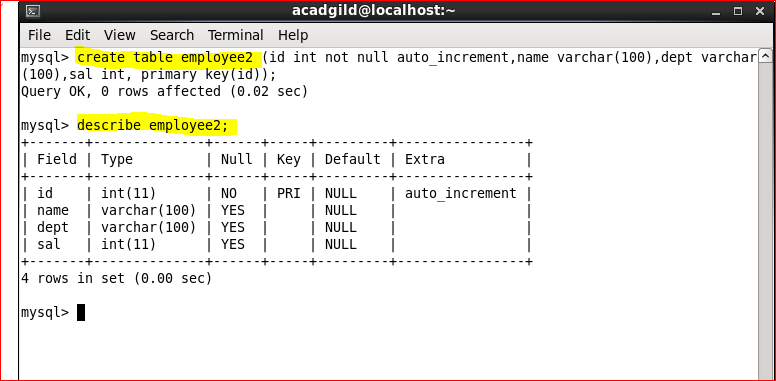
Here ‘-m’ specifies the number of map task that can be run simultaneously and ‘m1’ means that only one map task can run.

If we do not use **-m1**at the end of the statement, for each record in the MySQL table we will get separate files in the HDFS.

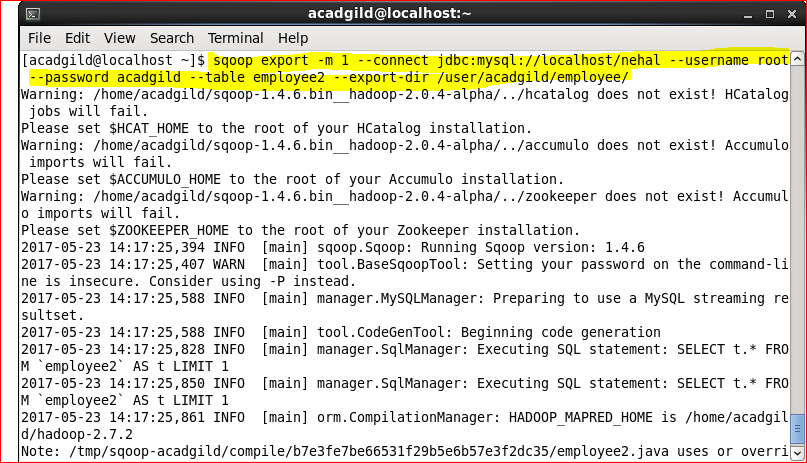
Now the data in RDBMS has been successfully imported into HDFS. By default, the files will be stored here: **/user/acadgild/employee/part-m-00000**file.

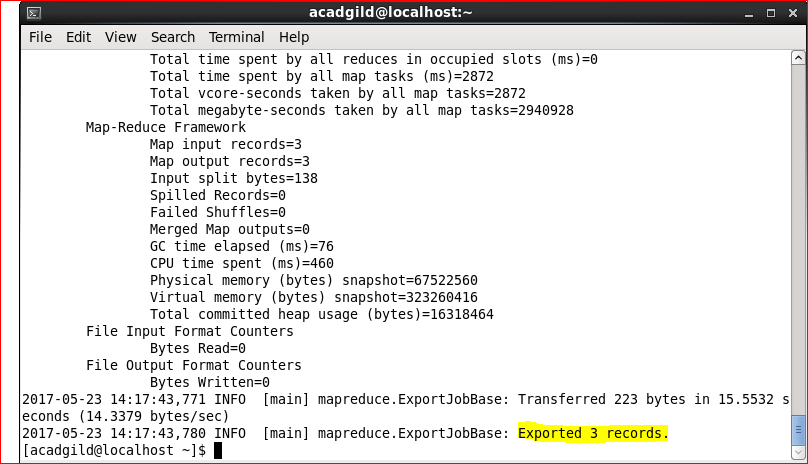


Create a table named employee2.



The files inside HDFS must have the same format as that of MySQL table, to enable the mapping of the data.





Sqoop calls the JDBC driver written in the –connect statement from the location where Sqoop is installed. The –username and –password options are used to authenticate the user and Sqoop, internally generates the same command against the MySQL instance.

The –table argument defines the MySQL table name that will receive the data from HDFS. This table must be created prior to running the export command. Sqoop uses the number of columns, their types, and the metadata of the table to validate the data inserted from the HDFS directory. When the export statement is executed, it initiates and creates INSERT statements in MySQL.

The -m argument sets the number of map jobs for reading the file splits from HDFS. Each mapper will have its own connection to the MySQL Server.

